

# The VVDS Data-Reduction Pipeline: Introducing VIPGI, the VIMOS Interactive Pipeline and Graphical Interface

M. SCODEGGIO,<sup>1</sup> P. FRANZETTI,<sup>1</sup> B. GARILLI,<sup>1</sup> A. ZANICHELLI,<sup>2</sup> S. PALTANI,<sup>3</sup> D. MACCAGNI,<sup>1</sup> D. BOTTINI,<sup>1</sup> V. LE BRUN,<sup>3</sup> T. CONTINI,<sup>4</sup> R. SCARAMELLA,<sup>2</sup> C. ADAMI,<sup>3</sup> S. BARDELLI,<sup>5</sup> E. ZUCCA,<sup>5</sup> L. TRESSE,<sup>3</sup> O. ILBERT,<sup>5</sup> S. FOUCAUD,<sup>1</sup> A. IOVINO,<sup>6</sup> R. MERIGHI,<sup>5</sup> G. ZAMORANI,<sup>5</sup> I. GAVIGNAUD,<sup>4,7</sup> D. RIZZO,<sup>4</sup> H. J. MCCracken,<sup>8,9</sup> O. LE FÈVRE,<sup>3</sup> J. P. PICAT,<sup>4</sup> G. VETTOLANI,<sup>2</sup> M. ARNABOLDI,<sup>10</sup> S. ARNOUTS,<sup>3</sup> M. BOLZONELLA,<sup>11</sup> A. CAPPI,<sup>5</sup> S. CHARLOT,<sup>8,12</sup> P. CILIEGI,<sup>5</sup> L. GUZZO,<sup>6</sup> B. MARANO,<sup>11</sup> C. MARINONI,<sup>6</sup> G. MATHEZ,<sup>4</sup> A. MAZURE,<sup>3</sup> B. MENEUX,<sup>3</sup> R. PELLÒ,<sup>4</sup> A. POLLO,<sup>6</sup> L. POZZETTI,<sup>5</sup> AND M. RADOVICH<sup>10</sup>

*Received 2004 December 23; accepted 2005 August 3; published 2005 October 21*

**ABSTRACT.** The VIMOS VLT Deep Survey (VVDS), designed to measure 150,000 galaxy redshifts, requires a dedicated data reduction and analysis pipeline to process in a timely fashion the large amount of spectroscopic data being produced. This requirement has led to the development of the VIMOS Interactive Pipeline and Graphical Interface (VIPGI), a new software package designed to simplify to a very high degree the task of reducing astronomical data obtained with VIMOS (Visible Multi-Object Spectrograph), the imaging spectrograph built by the VIMOS Consortium for the European Southern Observatory and mounted on Unit 3 (Melipal) of the VLT (Very Large Telescope) at Paranal Observatory (Chile). VIPGI provides the astronomer with specially designed VIMOS data-reduction functions, a VIMOS-centric data organizer, and dedicated data browsing and plotting tools, which can be used to verify the quality and accuracy of the various stages of the data-reduction process. The quality and accuracy of the data-reduction pipeline are comparable to those obtained using well-known IRAF tasks, but the speed of the data-reduction process is significantly increased, because of the dedicated nature of VIPGI. In this paper we discuss the details of the multiobject spectroscopy (MOS) data-reduction pipeline that has been implemented in VIPGI, as applied to the reduction of some 20,000 VVDS spectra, quantitatively assessing the accuracy of the various reduction steps. We also provide a more general overview of VIPGI capabilities, a tool that can be used for the reduction of any kind of VIMOS data.

## 1. INTRODUCTION

Over the last few years, the number of large telescopes available to the astronomical community has rapidly increased, as have the multiplexing capabilities of their instruments. While

a normal long-slit spectrograph on a 4 m class telescope could produce a few tens of spectra per night of observation, today a spectrograph such as VIMOS at the VLT, or the Two Degree Field (2dF) at the Anglo-Australian Telescope (AAT), can obtain several thousand spectra per night. This productivity increase has rendered obsolete traditional methods of data reduction and analysis, at least as long as these data must be reduced and analyzed in a timely fashion. It is clearly necessary to automate these operations as much as possible, in order to increase the speed with which they can be carried out, but without sacrificing the ability to analyze in detail the results of the various operations, and eventually to be able to manually intervene to change the way some of these operations are carried out. Moreover, it is necessary to develop an efficient and rigorous data organizer and archiver so that the available files and data are not lost among hundreds or thousands of similar data and files. General-purpose astronomical software packages are not well equipped for these tasks, as witnessed by the ad hoc data-reduction pipelines developed for most large spectroscopic surveys, in the case of both multislit (see, e.g., Colless et al. 1990 for the LDSS; Le Fèvre et al. 1995 for the CFRS) and multifiber observations (see, e.g., Colless et al. 2001 for the 2dF; Stoughton et al. 2002 for the SDSS).

<sup>1</sup> IASF Milano-INAf, Via Bassini 15, I-20133, Milan, Italy.

<sup>2</sup> IRA-INAf, Via Gobetti, 101, I-40129, Bologna, Italy.

<sup>3</sup> Laboratoire d'Astrophysique de Marseille, UMR 6110 CNRS-Université de Provence, BP8, 13376 Marseille, Cedex 12, France.

<sup>4</sup> Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées (UMR 5572), 14, Avenue E. Belin, F31400 Toulouse, France.

<sup>5</sup> INAF-Osservatorio Astronomico di Bologna, Via Ranzani, 1, I-40127, Bologna, Italy.

<sup>6</sup> INAF-Osservatorio Astronomico di Brera, Via Brera 28, Milan, Italy.

<sup>7</sup> European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany.

<sup>8</sup> Institut d'Astrophysique de Paris, UMR 7095, 98 bis Boulevard Arago, 75014 Paris, France.

<sup>9</sup> Observatoire de Paris, LERMA, 61 Avenue de l'Observatoire, 75014 Paris, France.

<sup>10</sup> INAF-Osservatorio Astronomico di Capodimonte, Via Moiriello 16, I-80131, Napoli, Italy.

<sup>11</sup> Università di Bologna, Dipartimento di Astronomia, Via Ranzani, 1, I-40127, Bologna, Italy.

<sup>12</sup> Max Planck Institut für Astrophysik, 85741, Garching, Germany.

In particular, recent surveys based on multifiber observations have been very successful in processing large amounts of spectroscopic data, but to their advantage is the fact that they have been observing relatively bright galaxies at low redshift. The relative brightness of the observed galaxies makes the sky subtraction process a noncritical one, whereby the determination of the sky spectrum required only a few percent accuracy level. Significantly more challenging is the data-processing task for a survey extending to higher redshift than that covered by the 2dF (Colless et al. 2001) and SDSS (Strauss et al. 2002) surveys. In this case it is necessary to derive spectra of very faint galaxies, with a surface brightness of only a few percent of that of the sky, and in the red part of the optical spectrum ( $\lambda \sim 9500 \text{ \AA}$ ), a wavelength domain contaminated by many strong sky emission features. Early work in this challenging domain was provided by the Canada-France Redshift Survey team (Le Fèvre et al. 1995), but nowadays the high data throughput of 8 m class telescopes require significant research and development efforts to obtain a spectroscopic data-reduction pipeline that is capable of properly handling the data these telescopes are obtaining.

Among the last-generation spectrographs, VIMOS is perhaps the most challenging in terms of data production, since it currently has the highest multiplexing capabilities available to any astronomer. It has been specially designed to be a survey instrument, and therefore its multiplexing capabilities have been pushed to the maximum: during a single exposure, up to 1000 spectra can be obtained in multiobject spectroscopy (MOS) mode (6400 using the Integral Field Unit [IFU] mode). Difficulties for the user in the data-reduction process begin when trying to find one's way among the large number of raw science and calibration files, and they increase further along the reduction process, due to the number of calibrations and corrections that have to be applied to the data. To give an example, even just a single pointing in MOS mode, with its minimal set of calibration data (five bias, five flat field, and one arc lamp exposure), is equivalent to 48 FITS files being produced (raw data only).

For this reason, we were contracted by the European Southern Observatory (ESO) to deliver all the elements necessary to build a VIMOS-specific automatic data-reduction pipeline. In collaboration with the ESO Data Management Division, we have developed a C library of data-reduction procedures, the VIMOS Data Reduction Software (DRS; see Scodeggio et al. 2000), which is now being used for the online reduction of VIMOS data at ESO. Because of its online use, the ESO pipeline must work in a completely automatic fashion; but while this is an optimal choice for obtaining quick quality assessments for VIMOS observations, it is a far less acceptable one for a careful and complete science data-reduction pipeline. Faced with the task of reducing some 100,000 spectra observed as part of the VLT VIMOS Deep Survey (VVDS; see Le Fèvre et al. [2005] for a detailed description of the survey), we have used the DRS to produce a semiautomatic data-reduction pipe-

line, the VIMOS Interactive Pipeline and Graphical Interface (VIPGI). With VIPGI we have kept the DRS's ability to perform very fast data-reduction processes, but we also have added many data-reduction quality control points, a user-friendly graphical interface, and a simple but effective method to organize the raw and reduced data.

VIPGI was specifically designed to carry out the reduction of the VVDS data, and its performance in this task is the focus of this paper. But VIPGI's capabilities are far more general. It can be used to reduce VIMOS data of any kind, and not just MOS data. In particular, it can be used to reduce VIMOS IFU data, a task that is far more complex than the relatively simpler reduction of MOS data, and that is described in detail in a separate paper (Zanichelli et al. 2005, hereafter ZAN05). VIPGI capabilities are general enough, at least as far as MOS data are involved, that it could be used as a generic pipeline tool for the reduction of MOS data from any spectrograph, and work is already under way to develop this functionality (see Paiono et al. 2004). So far, VIPGI has been extensively tested while reducing more than 20,000 MOS spectra for the VVDS, and a smaller number of MOS and IFU spectra from other, independent projects. It has been publicly available to the astronomical community since 2005 June, but it can also be used in either Milan, Italy, or Marseille, France, by members of the ESO community under the supervision of a VIRMOS Consortium astronomer, who can provide guidance on how to handle nonstandard situations.<sup>13</sup>

In this paper we briefly present the characteristics of the VIMOS spectrograph (§ 2), discuss some of VIPGI's general concepts (§§ 3 and 4), and then focus on the reduction of VVDS MOS data (§ 5) and the quality of the final VVDS spectra (§ 6). The data organizer and the data browsing and plotting facilities are briefly described in § 7.

## 2. THE VIMOS SPECTROGRAPH

VIMOS is an imaging spectrograph installed on Unit 3 (Melipal) of the ESO Very Large Telescope (VLT) at the Paranal Observatory in Chile (see Tables 1 and 2; see also Le Fèvre et al. [2000, 2002] for a detailed description of the instrument and its capabilities). The driving design concept for the instrument is to cover as much of the unvignetted part of the focal plane as possible at the VLT Nasmyth focus (a circular area with a diameter of  $22'$  on the sky). Since this large area corresponds to a very large linear scale (almost 1 m), it was decided that coverage would be achieved by splitting the instrument into four identical optical channels arranged next to each other and supported by the same mechanical structure. Each optical channel is a classical focal reducer imaging spectrograph, with a collimator, a parallel beam where the dispersive element (a grism) is inserted, and a camera that focuses the beam onto a  $2048 \times 4096$ ,  $15 \mu\text{m}$  pixel EEV CCD. The focal

<sup>13</sup> See <http://cosmos.mi.iasf.cnr.it/pandora/vipgi.html>.

TABLE 1  
GENERAL CHARACTERISTICS OF THE VIMOS SPECTROGRAPH

Parameter	Value
Field of view .....	$4 \times 7' \times 8'$
Detector area .....	$2048 \times 4096$ pixels
Scale .....	$0''.205 \text{ pixel}^{-1}$
Direct imaging filters .....	$U'BVRIZ$
Wavelength range .....	370–1000 nm
Slit masks .....	Four masks, each $\approx 300 \times 300$ mm, laser cut
Multiplex .....	$\approx 800$ simultaneous slits, each $10''$ long at $R \approx 200$
Slit sampling .....	4.88 pixels for $1''$ slit
Slit width and length .....	Larger than $0''.35$ , any length and shape

plane is flattened by a field lens at the instrument entrance, to allow for flat multislit masks, and a folding mirror is inserted into the collimator, to fold the beam and to reduce the instrument's overall size. The field of view covered by each channel (generally referred to as a VIMOS quadrant) is approximately  $7' \times 8'$ , with a pixel scale of  $0''.205 \text{ pixel}^{-1}$ . An approximately  $2'$  wide gap is present between quadrants (see Fig. 1).

A passive flexure compensation mechanism is in place, mounted on support struts attached to the folding mirror supports, and has been optimized to achieve a reasonable compromise at every instrument position while it rotates with the telescope Nasmyth rotator. With this system in place, the effect of flexure is limited to an image motion of less than  $\pm 2$  pixels on the CCD. Such image stability is very important to permit a satisfactory removal of fringing effects, which are very pronounced above  $8500 \text{ \AA}$ , because of the thinned, back-illuminated CCDs used in VIMOS. Also, because the flexures affect the four optical channels differently, the flexure compensation mechanism is vital to ensure a satisfactory centering of astronomical objects within MOS slits in all four quadrants contemporaneously.

For spectroscopic observations, six different grisms provide a spectral resolution ranging from  $R \sim 200$  to 2500 (see Table 2). Order-sorting filters are used to avoid an overlap between first and second grating orders. Laser-cut masks (one per quadrant) are used for MOS observations. The number of slits that can be placed on each mask varies from approximately 40 at high spectral resolution up to approximately 250 at low spectral resolution. An imaging exposure is required as the starting point of the mask design and cutting process (see Bottini et al. 2005).

It is also possible to carry out integral field spectroscopic observations using the IFU. The MOS and IFU modes entirely share the four VIMOS optical channels, but in IFU mode the light collected by the IFU head (located on a side of the focal plane assembly; see Fig. 1) is transferred to the four channels via optical fibers that connect the IFU head to four so-called IFU masks, which are special support structures that are placed in the instrument focal plane instead of the laser-cut MOS masks. For more details on the VIMOS IFU mode, see ZAN05.

TABLE 2  
IMAGING CHARACTERISTICS OF THE VIMOS SPECTROGRAPH

Grism	Wavelength		Dispersion	Multiplex (rows)
	Range ( $\lambda$ )	Resolution ( $\text{\AA pixel}^{-1}$ )		
LR blue .....	370–670	180	5.3	4
LR red .....	550–950	210	7.3	4
MR .....	500–1000	580	2.5	2
HR blue .....	415–620	2050	0.51	1
HR orange .....	520–760	2150	0.6	1
HR red .....	630–870	2500	0.6	1

NOTE.—Values for spectral resolution are given for  $1''$  slit.

### 3. THE VIMOS INTERACTIVE PIPELINE AND GRAPHICAL INTERFACE

The core of the VIPGI pipeline is a set of routines that perform the data reduction, developed in collaboration with the ESO Data Management Division, following their general requirements and constraints for a VLT instrument data-reduction pipeline. Since those constraints prevented us from using IRAF<sup>14</sup> tasks as the building blocks for the data-reduction routines, these were coded using the C programming language to obtain the maximum speed for such computationally intensive tasks. Routines range from the basic opening of a file and reading of its content to more complex wavelength calibration, spectral extraction, and IFU three-dimensional data cube reconstruction, and they are organized in a relatively small number of reduction recipes. Generic tasks such as the handling of FITS files or the World Coordinate System, and the detection of stars within imaging exposures (for photometric or astrometric calibration purposes), are devoted to special-purpose external software packages (the CFITSIO and WCSTools libraries and SExtractor, respectively; see Pence 1998; Mink 1998; Bertin & Arnouts 1996) that have been included within VIPGI.

The global data-reduction scheme implemented by VIPGI is a fairly traditional one, broadly following the one implemented by the IRAF long-slit package. However, a special effort was made to achieve a very high efficiency during the repeated application of this scheme to the large set of VVDS data by tailoring all aspects of the data-reduction scheme to the specific characteristics of VIMOS. Still, the various reduction functions are general enough that it could be adapted for the reduction of data produced by any MOS spectrograph, with a minimal effort. Reduction steps that are normally always executed in the same sequence are always grouped into a single recipe (e.g., bias subtraction, flat-field correction, and bad-pixel cleaning), and all slits in a given mask are always reduced concurrently so that the complete reduction of one VIMOS MOS scientific

<sup>14</sup> IRAF (Image Reduction and Analysis Facility) is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

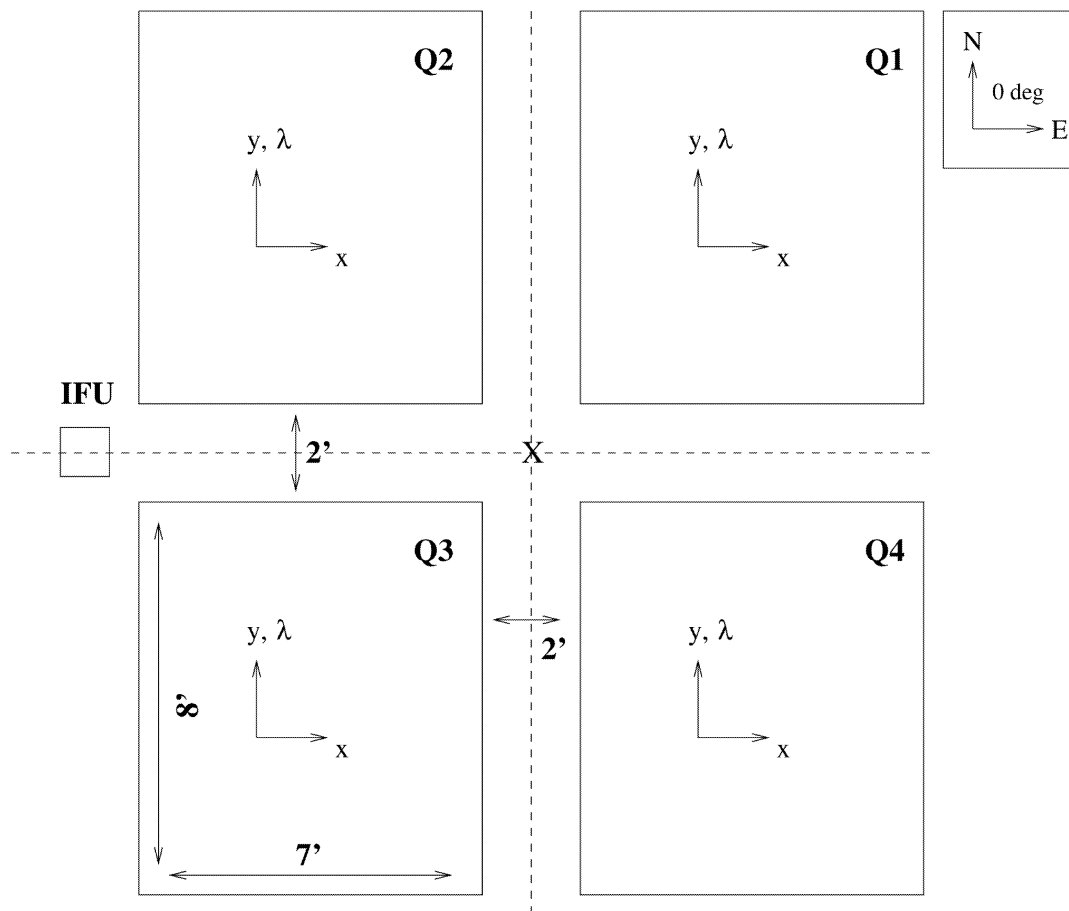


FIG. 1.—Schematic representation of the VIMOS focal plane projection on the sky.

exposure can be completed in just four basic steps. However, there is no single “do it all” recipe that can be fed a bunch of raw data frames to produce completely reduced images or spectra, as we assume that the astronomers will need and want to check at least some of the intermediate data-reduction steps. To help astronomers keep the details of the data-reduction process under control, the detailed behavior of each recipe can be customized via a set of input parameters that are stored in a parameter file.

All recipes have been written to work with files in FITS format. To avoid increasing the already large number of files, the different mid-reduction products, together with the various calibration tables needed for the reduction process, are “appended” as extensions to the original FITS file, instead of creating independent files. As a consequence, the results of the reduction of one spectroscopic exposure are contained within only four files (one per quadrant), each containing up to 12 FITS image and binary table extensions. Thus, these files can easily be over 100 MB in size.

The C recipes automate to a very large extent the task of

reducing VIMOS data, but they do not address two important and problematic areas of the global data-reduction activity: the need to organize a large volume of VIMOS data, and the need for quick and easy browsing of the data at the various stages of data reduction. Without a satisfactory solution to these two problems, it is impossible to obtain an efficient data-reduction pipeline. It is mainly to address these two needs that we have built the VIPGI graphical user interface, using the programming language Python. This high-level “scripting” language is becoming one of the industry standards for these kind of tasks, as witnessed for example by the recent distribution of a new Python-based IRAF interface (PyRAF; see Greenfield & White 2000). VIPGI provides a VIMOS-specific data organizer, designed to help the user select the correct data to use at each stage of the data-reduction process, in addition to a simple interface to the underlying C reduction procedures, tools that allow simple and quick browsing through the VIMOS data, and powerful plotting tools to view and analyze one- and two-dimensional extracted spectra. No built-in image display tool is provided, but users can configure VIPGI to use their preferred

viewer. VIPGI features are presented in the following sections. Additional information and screen shots can be found on the Web to better understand all of the VIPGI features.<sup>15</sup>

In total, the software that was developed specifically for this project is composed of ~151,000 lines of C code and 16,000 lines of Python code. The overall performance of the VIPGI pipeline of course depends on the computer hardware being used (most importantly, on the amount of available physical memory, as the pipeline recipes do not use any advanced memory management system) and on the kind of data being reduced. As an example, the reduction of a typical VVDS deep pointing (four masks for a total of ~550 slits, 10 jittered exposures to be combined) on a Linux PC with 1 GB of physical memory takes approximately 25 CPU seconds for a complete wavelength calibration of the four quadrants, and 1170 CPU seconds to complete the reduction of the science data. From the end-user point of view, it takes approximately 1 hr to go from raw data to a set of fully reduced spectra, including the CPU time required to run the reduction recipes, the time spent selecting the appropriate input files for those recipes, and the visual inspection and verification of the instrument calibrations described below.

#### 4. THE INSTRUMENT CALIBRATIONS

A large increase in the efficiency of the VIPGI data-reduction process is achieved by simplifying to the maximum the handling of calibrations for VIMOS data. VIPGI recipes require a preexisting calibration of the instrument properties as part of their input, which is used as a first guess that will be subsequently refined by the recipes themselves; such a calibration is always provided to the astronomer, stored in the raw data FITS header. This operational choice saves the astronomer the time normally devoted to the task of deriving a first-guess calibration. It also results in a more homogeneous data-reduction process, since all astronomers effectively start the reduction from the same calibration data, an important fact for a survey instrument like VIMOS. A fundamental component of the whole pipeline architecture is therefore the VIMOS instrument model, which analytically describes the main calibration relations required for the extraction of object spectra from VIMOS spectroscopic observations. To achieve its large multiplexing factor, VIMOS produces spectroscopic data in which to four different spectra are present along a given CCD column (contrary to most other spectrographs, in which only a single spectrum is present along a CCD row or column). As a consequence, the instrument model has to be a fully bidimensional mapping of the VIMOS focal plane. The model is separated into three different components:

1. *Optical Distortion Model*.—Provides a mapping between positions on the VIMOS focal plane and pixel coordinates on the CCD frame. This mapping is obtained for an arbitrarily

fixed wavelength and is described by two independent polynomial relations. Since VIMOS spectra are dispersed along the CCD columns, the relation that gives the  $x$ -pixel coordinate as a function of position in the focal plane is wavelength independent and is influenced only by the optical and mechanical layout of the instrument. The relation that gives the  $y$ -pixel coordinate, however, is also dependent on the choice of the reference wavelength.

2. *Curvature Model*.—Provides a description of the geometrical shape of each spectrum on the CCD, to allow for its tracing and extraction. As optical distortions in the VIMOS focal plane are very small, this model is obtained using a low-order ( $\leq 2$ ) polynomial.

3. *Inverse Dispersion Solution*.—Provides the mapping between wavelength and pixel coordinates along the geometrical shape traced by the curvature model, and measured as offsets with respect to the reference wavelength and pixel position defined by the optical distortion model. In addition, this mapping is described by a relatively low-order polynomial, with the order depending on the grism being used.

The instrument model can be used at two different levels of detail to obtain a description of VIMOS data. The local model provides the highest level of detail, as it describes each MOS slit or IFU fiber spectrum individually. The global model instead provides a global description for one whole CCD frame, which is significantly more robust than that provided by the local model, as it is derived using many tens (MOS case) or hundreds (IFU case) of spectra, although at a reduced level of detail. The optical distortion model, by its very nature, exists only at the global level.

The starting point for each recipe is always a global model whose parameters are periodically derived by VIMOS instrument scientists using data taken with specially designed calibration masks, which are stored in the raw data FITS file header when the observations are carried out. This model is used to derive a local model for each spectrum, which is then refined by fitting its parameters to the real data. Finally, an updated global model is obtained by fitting the parameters of all available local models. Both the global and local models are stored within the data file itself, to be available for all reduction recipes. In this way, the best possible calibration is obtained for each individual VIMOS exposure. Still, with accurately calibrated global and local models, it is important to decide which one to use for the reduction of scientific observations. The choice actually used within VIPGI is described in detail in the following sections.

Since the four VIMOS quadrants correspond to four physically distinct cameras within the instrument, each quadrant is characterized by its own instrument model. Therefore, all calibration and science data-reduction procedures are carried out on the data from each quadrant independently, and only at the end of the data-reduction process, and only in some special

<sup>15</sup> See <http://cosmos.mi.iasf.cnr.it/pandora/vipgi.html>.

cases (such as IFU data cube reconstruction), are the data from all four quadrants brought together into a single data set.

#### 4.1. The Location of Spectra on the CCD Frame

The optical distortion and curvature models are derived from the accurate determination of spectral locations on the CCD frame. A halogen lamp flat-field exposure or a scientific exposure (provided this exposure is long enough to supply enough signal in the background continuum, typically 200–300 counts) are used for this task. Using MOS slit positions that are stored in the raw data FITS file header (or IFU fiber end positions, which, being fixed, are stored in a permanent calibration table), and the global instrument model parameters (also stored in the FITS header), a first guess for the position of each slit or fiber spectrum on the CCD is obtained. Starting from this position, a search is made for the edges of the illuminated area created by the spectrum on the CCD. The shape of these edges is fitted with a polynomial to provide the updated local determination of the curvature model and a very precise determination of the spectrum position on the CCD. Curvature model parameters and position measurements for all slits are then fitted with polynomials to provide the updated global determination for the optical distortion and curvature model.

MOS spectra are generally located on the CCD with a typical uncertainty of 0.2 pixels, provided that the mask layout automatically derived by the VIMOS Mask Preparation Software (VMMP; see Bottini et al. 2005) is not altered by the astronomer by manually adding slits on the mask. VMMP is in fact enforcing the presence of a minimum gap between adjacent MOS slits in order to make the measurement of spectral edges possible. IFU fiber spectra are not so cleanly separated, as there are no sharp slit edges to locate, and for this reason a slightly different approach is followed in that case (see ZAN05 for details).

With the exception of IFU data taken with high-resolution grisms, the accuracy with which the instrument's global and local models can describe the location of spectra on the CCD is comparable, and therefore by default VIPGI uses the more robust global model to carry out this description. For high spectral resolution IFU data, it is basically impossible to obtain a reliable global model, because of the peculiar geometry with which the illuminated fibers are distributed on the CCD (see ZAN05), and therefore the more fragile local model must be used. VIPGI data browsing facilities make it possible to carry out an accurate visual check of the quality of spectral location by displaying a raw VIMOS data frame upon which is superposed the reconstructed location of each spectrum.

#### 4.2. The Wavelength Calibration

The inverse dispersion solution, which provides the wavelength calibration for VIMOS spectra, is obtained by measuring the position on the CCD of a number of strong emission lines. This operation can be carried out using either a calibration

lamp exposure (a combination of He, Ne, and Ar lamps can be used for this purpose) or the night-sky lines present in a science exposure. The list of lines to be measured is provided by a predefined line catalog, while the location of the spectra is described by the previously derived instrument model, using the local model obtained for each slit or fiber spectrum. The two-dimensional spectrum produced by a MOS slit is separated into one-dimensional spectra, one for each pixel spanned by the length of the slit (for IFU fiber spectra, only the central pixel spectrum is extracted). Within these spectra, the position of lamp lines is measured, computing their baricenter within a predefined extraction window, and the measured positions are fitted against the known line wavelengths, using a low-order polynomial function and an iterative  $\sigma$ -clipping procedure. This fit provides the coefficients for the local model inverse dispersion solution for each slit.

The accuracy of the wavelength calibration changes slightly from grism to grism, but the rms residuals around the best-fitting relation typically amount to one-fifth of a pixel when the calibration is carried out using a lamp exposure. The accuracy that can be achieved using night-sky lines is typically 10 times worse, due to the combination of a weaker signal for each line with lines crowding in the red part of the spectrum, or with a sheer lack of lines in the blue part. For the VVDS data obtained with the low-resolution red grism, which produces spectra with a linear dispersion of  $7.14 \text{ \AA pixel}^{-1}$ , a fit with a third-degree polynomial to calibration lamp exposures results in a wavelength calibration with a median rms residual of  $1.005 \text{ \AA}$  (measured over a total of 16,936 slits). Although the distribution shows the presence of a tail extending to high rms residual values, only approximately 0.9% of the slits show a wavelength calibration with rms residuals larger than  $3.0 \text{ \AA}$ . No systematic variation in the calibration accuracy is observed as a function of slit position within the MOS masks, or as a function of the VIMOS optical quadrant.

For wavelength calibration, unlike spectra location, the local instrument model always provides the most accurate calibration information, and therefore it is the one used in the reduction of scientific observations. Using VIPGI data browsing and plotting tools, it is possible to carry out an accurate visual check of the quality of the wavelength calibration, and also recompute such calibration for individual slits or fibers after having eliminated significantly deviant spectral lines from the list.

The accuracy and stability of the wavelength calibrations are obviously of greatest importance for spectroscopic surveys, which are the most natural observing program for VIMOS. However, the accuracy with which the inverse dispersion solution maps the true wavelength calibration of VIMOS observations cannot be quantified solely on the basis of the rms residual values obtained in the calibration procedure. In fact, changes in temperature or flexures within the instrument might significantly alter this mapping between the time when the science and calibration exposures are taken (calibration lamp exposures are typically taken only at the end of a sequence of

exposures, and sometimes during the following day). Therefore, it is important to reconsider the issue of wavelength calibration accuracy on the basis of the reduced science data and redshift measurements, and we do that in § 6.

## 5. SCIENCE DATA REDUCTION

The first step in the reduction of VIMOS science data is the canonical preliminary reduction of CCD frames, which includes prescan level and average bias frame subtraction, trimming of the frame to eliminate prescan and overscan areas, interpolation to remove bad CCD pixels, and flat fielding. For VVDS data taken with the low-resolution red VIMOS grism, a flat-fielding of the data is not carried out, since flat-field exposures, which are obtained using an internal halogen lamp that illuminates the reflective cover of the telescope Nasmyth focus shutter screen, show the presence of a significant amount of fringing. The spatial frequency of this fringing pattern is too high to be reliably removed with a surface-fitting procedure, and the pattern present in flat-field exposures is completely different from the one present in science exposures, due to the different spectral energy distribution of the halogen lamp and of the night sky. As the noise introduced by flat-fielding VVDS data with such fringing-rich flat-field frames is actually larger and more structured than that produced by the CCD pixel-to-pixel sensitivity variations, it was decided to eliminate the flat-fielding altogether from the VVDS data-reduction scheme. This choice introduces some uncertainty in the spectrophotometric calibration of VVDS spectra. As the rms in the pixel-to-pixel sensitivity variations is measured to be approximately 2%, and the average spectrum is extracted by integrating the object signal over 11 pixels, the expected contribution from this lack of flat-fielding to the uncertainty of the flux calibration of VVDS spectra is of the order of 0.6%, significantly smaller than that introduced by slit losses and varying sky transparency (estimated to contribute uncertainties at the 20% level).

After the preliminary reduction step, subsequent data-reduction steps are carried out on all MOS slits individually, one slit at a time. For each individual spectroscopic exposure, the presence of flexure within the instrument imposes a readjustment, based on the science data themselves, of the wavelength calibration derived from an arc lamp exposure. This operation is carried out using a number of sky lines by comparing their known wavelength with the one derived using the local model inversion dispersion solution for the given slit. If any discrepancy is measured, an offset is introduced into the inverse dispersion solution to compensate for it. In practice, this is a virtual offset that is to be applied to all pixel coordinates of a given slit or fiber spectrum at the moment of its extraction (discussed below) to counteract the offset introduced by the instrument flexures.

The following steps in the data-reduction procedure include object detection and sky subtraction from each MOS slit spectrum within each individual MOS science exposure. The object

detection procedure is based purely on the data themselves and does not use any prior information from the MOS mask design procedure to locate object spectra within the MOS slits. The raw-data slit spectrum is collapsed over a user-defined wavelength interval, following the geometrical shape defined by the curvature model for the given slit, to produce a slit cross-dispersion profile. A robust determination of the average signal level and its rms variations in this profile is obtained using an iterative  $\sigma$ -clipping procedure, and objects are detected as groups of contiguous pixels above a given detection threshold, expressed in units of the signal level rms variations. For VVDS data, we use a detection threshold of  $2\sigma$  and a minimum object size of 3 pixels. Possible spurious detections, often taking place close to the slit edges, are removed at a later stage, during the visual inspection of the extracted spectra. The sky-subtraction process reproduces the standard long-slit background-subtraction procedure. For each slit, a median estimate of the sky spectrum is obtained using all the pixels that are devoid of object signal (on the basis of the object detection results just obtained), and this sky spectrum is then subtracted from the total slit data. The sky background must be estimated separately for each slit, because of the varying contributions produced by zeroth and second grating orders from other slits along the same CCD columns. It must also be estimated separately for each individual science exposure, because of the significant variations in OH line strength over the timescale of a typical spectroscopic exposure (15 to 45 minutes). However, in the case of tilted slits, in which lines of constant wavelength produce spectra on the CCD that are tilted compared to the CCD rows, this same procedure is postponed until after the two-dimensional extraction and wavelength rectification step described below.

The sky-subtracted slit spectra are then two-dimensionally extracted using the tracing provided by the slit curvature model, and resampled to a common linear wavelength scale. Only after this point are the single exposures of a jitter sequence combined together. The main reason for using a jitter sequence for the VVDS observations is to help eliminate the strong fringing pattern present in VIMOS observations obtained with the low-resolution red grism. This pattern is relatively constant for the duration of a typical jitter sequence observation (a few hours), while both the overall intensity and the relative strength of the individual emission lines within the sky background can have very significant variations over those timescales. It was therefore decided that for the implementation of the VIPGI jitter sequence reduction procedure, the subtraction of this relatively constant fringing pattern should be postponed until after the removal of the more rapidly varying sky background.

All VVDS observations are obtained using a five-offset pattern, with offsets of  $0''.75$  between exposures. This specific choice for the jitter parameters is primarily dictated by two considerations. First, a pattern with many offset positions results in a more accurate sky subtraction in the blue part of the spectrum (where the sky signal is rather low) with respect to

a simple two-position (traditionally referred to as ABBA) pattern. Second, the product of the number of offset positions and the offset width basically dictates the length of the slits, and this in turn dictates the number of spectra that can be obtained in a single VIMOS pointing. A large offset width (say,  $2''$ ), which would be desirable in order to increase the overall accuracy of the sky subtraction, would result in a relatively small number of spectra per pointing when coupled with a large number of offset positions, thus severely increasing the telescope time requirements for the VVDS. As a compromise between these irreconcilable demands, we have fixed the number of offset positions to five, thus choosing an offset width of  $0''.75$ , comparable to the typical size of the seeing disk at the VLT.

To obtain a completely automated fringing pattern removal, the  $N$  exposures that are part of a jitter sequence are combined together twice. First, the  $N$  two-dimensionally extracted spectra for each slit are median-combined, without taking into account the jitter offsets, to produce a two-dimensional sky subtraction and fringing pattern residual map. Since the jitter offsets are somewhat smaller than the typical object size, the median combination is not enough to completely eliminate the object spectrum from these residual maps, and therefore the object spectra areas are masked out from the single-slit spectra before the combination, interpolating across adjacent sky-only parts of the spectrum. The residual map is then subtracted from all the  $N$  two-dimensional single-exposure slit spectra. At this point a second combination is carried out, this time taking into account the jitter offsets among the  $N$  individual two-dimensional slit spectra: object positions within these spectra, measured as part of the object detection process described above, are used to derive a very accurate determination of the telescope offsets applied through the jitter sequence. The single-exposure residual-map-subtracted spectra are offset to compensate for the effect of the jitter, and a final average two-dimensional spectrum for each slit is obtained.

Figure 2 shows two examples of such two-dimensional spectra obtained with and without the subtraction of the residuals map, to illustrate the effectiveness with which the adopted procedure removes the fringing pattern from the data. This removal is most efficient when only a faint-object spectrum is present in the slit, as in this case, in which the adopted jitter pattern allows the removal of the object spectrum from the no-offset combination of the jittered exposures without any problems. As the objects get brighter and larger, their extended spectra can prevent a reasonable estimate of the residual pattern, and therefore the removal of the fringing residuals becomes less accurate.

The object detection process is repeated on the combined two-dimensional spectra to produce the final catalog of detected spectra, and a one-dimensional spectrum is extracted for each detected object, using the Horne optimal extraction procedure (Horne 1986). Finally, spectra are flux calibrated using a simple polynomial fit to the instrument response curve derived from observations of spectrophotometric standard stars, and cor-

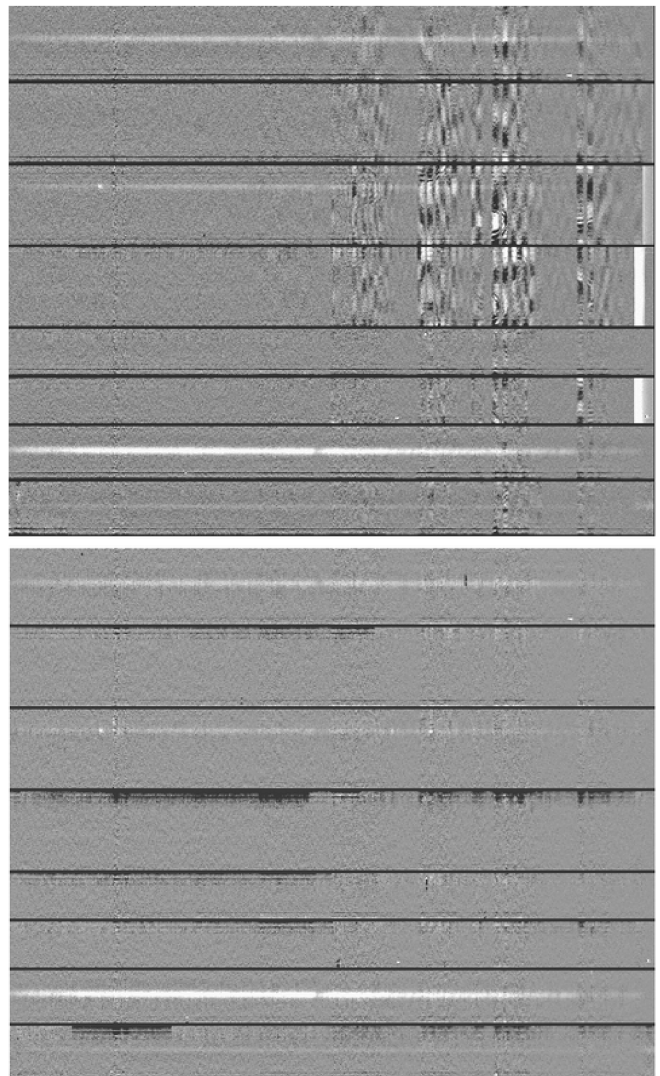


FIG. 2.—Example of the effectiveness of the fringing subtraction procedure. *Top*: Results of combining 10 jittered VVDS exposures without the subtraction of the fringing residuals. *Bottom*: Results of combining the same 10 exposures with the subtraction. Only a small portion of the entire CCD frame is shown, with wavelength increasing from left to right. Some background-subtraction artifacts that are related to the fringing subtraction procedure are visible. On the top half there are some grism zeroth-order residuals visible at the redmost extreme of the spectra (positive residuals at constant wavelength), which are subtracted off by the fringing correction procedure. This procedure introduces some residuals itself at the slit edges (in this case, negative residuals), where the signal for the fringing residuals map is contributed by just one exposure in the jitter sequence.

rected for telluric absorption features. The last correction is based on a template absorption spectrum derived for each combined jitter sequence from the data themselves. Figure 3 shows a block diagram summary of the various steps involved in the reduction of a MOS observation with VIPGI.



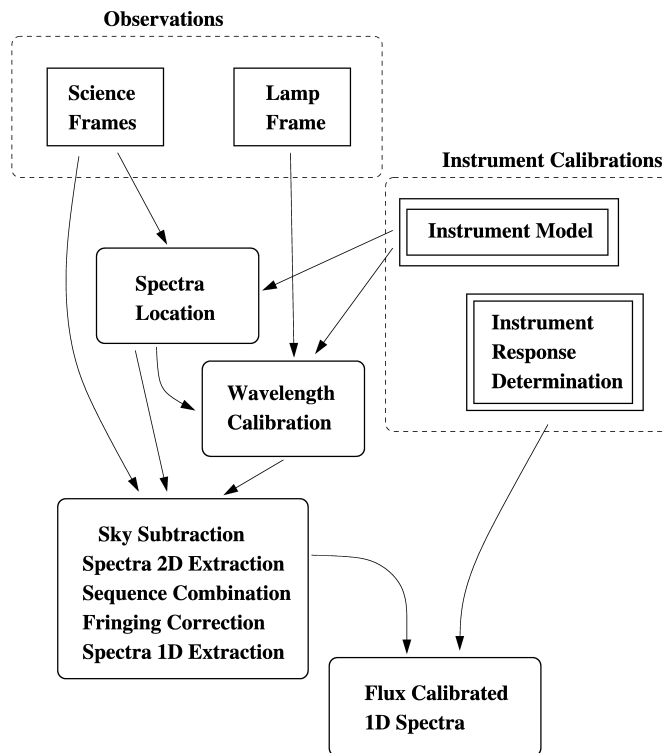


FIG. 3.—Block diagram summarizing the main steps involved in the reduction of MOS data using VIPGI.

## 6. THE QUALITY OF SPECTRAL EXTRACTION AND WAVELENGTH CALIBRATION

One rather obvious criterion we have used to assess the quality of our data-reduction pipeline is to compare its results with those obtained in a very careful data reduction carried out using standard and well-tested IRAF tasks. No ad hoc IRAF-based scripts were written for this comparison, both because of time constraints and to avoid introducing into the comparison the uncertainty of possible coding errors. As a result, the manual IRAF reduction was quite time consuming, and the comparison was carried out only for a couple hundred spectra. It has shown nonetheless that the quality of VIPGI-reduced spectra, in terms of continuum shape and signal-to-noise ratio (S/N) at all wavelengths, is basically the same as the one obtained with a standard IRAF reduction.

More systematically, we have compared the  $R$  and  $I$ -band total magnitudes from the photometric catalog used to build the VVDS sample with the magnitude derived by integrating the flux in the observed spectrum with the appropriate filter response curve. The results of such a comparison are shown in Figure 5 of Le Fèvre et al. (2005) for the  $I$ -band magnitudes, and in Figure 4 for the  $R$ -band magnitudes. There is a satisfactory agreement between the spectral fluxes and the photometric magnitudes, although a significant scatter can be seen in this comparison. A significant fraction of this scatter is cer-

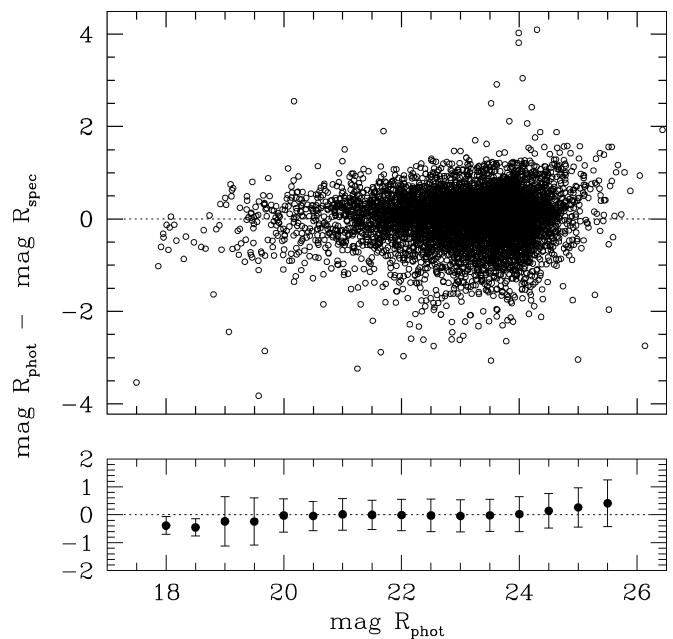


FIG. 4.—Comparison between the  $R$ -band magnitudes from the photometric VVDS catalog and the magnitudes computed by integrating the observed spectrum with the equivalent filter response curve. *Top*: Full distribution of data points; *bottom*: mean and 1  $\sigma$  rms scatter, in half-magnitude bins.

tainly due to the flux-calibration procedure adopted for the VVDS: we are not trying to obtain a nightly flux calibration, but are only using a generic calibration template to reproduce the correct spectral energy distribution shape, without any specific attempt to obtain the right absolute flux normalization. Variations in sky transparency or seeing conditions are not taken into account by our flux-calibration procedure and thus show up in the scatter visible in the figure. Another contribution to the scatter comes from slit losses, the fraction of the total object flux that is not transmitted through the slit because of both random errors in the centering of objects within the MOS slits and variations in the ratio between object and slit size (the latter being fixed to  $1''$ ).

The most important quality checks on VVDS data concern wavelength calibration and the accuracy and stability of redshift measurements we can obtain with VVDS data reduced with VIPGI. As discussed already in § 4.2, the accuracy of the wavelength calibration relation measured on the lamp exposures does not provide a complete picture of the global calibration accuracy for VIMOS data. We therefore use the reduced spectra themselves to assess this quality.

Using redshift measurements, we can test the absolute stability of the wavelength calibration. This test can be carried out in two different ways: within VIPGI, by comparing repeated observations of the same objects, and via an external comparison, by comparing VIMOS VVDS observations reduced with VIPGI with FORS2 observations obtained by the K20 and

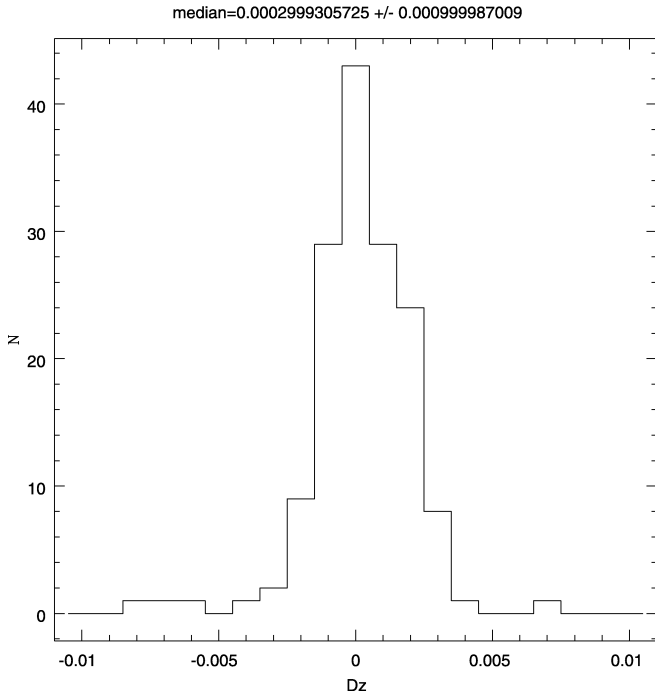


FIG. 5.—Distribution of differences in redshift measurements from repeated VVDS observations of a set of 150 objects.

GOODS surveys in the CDFS. As we are interested in measuring the stability of the wavelength calibration, and not the overall reliability of the redshift measurement process, in this comparison we use only those spectra for which the separate redshift measurements are in agreement (i.e., within a  $z$  interval of 0.1). Within the VVDS, we have 150 repeated observations that produce redshift measurements in agreement. The distribution of these differences is shown in Figure 5: the differences have an average value of 0.0003, with an rms scatter of 0.0010. We also have 41 and 27 galaxies in common with the K20 and GOODS survey, respectively, with redshift measurements in agreement. The differences have an average value of  $-0.0004$  and  $-0.0006$ , with an rms scatter of 0.0018 and 0.0013, respectively. The global distribution of these differences is shown in Figure 6. Taking account the uncertainties due to the cross-correlation or line location procedures themselves, these measured differences show that the absolute stability of the wavelength calibration in the VVDS data reduced with VIPGI is at the level of 2 to 3 Å, significantly less than the linear dispersion coefficient for these spectra ( $7.14 \text{ Å pixel}^{-1}$ ).

One final measure of the quality of the spectra extraction process is the fraction of failures in this process. There are many possible sources for these failures, from spurious detections in the parent photometric catalog, which result in no object being present in the slits, or errors in the astrometric calibration of the catalog, resulting in a poor centering of the object in the slits, to slits vignettted by the telescope guide camera probe,

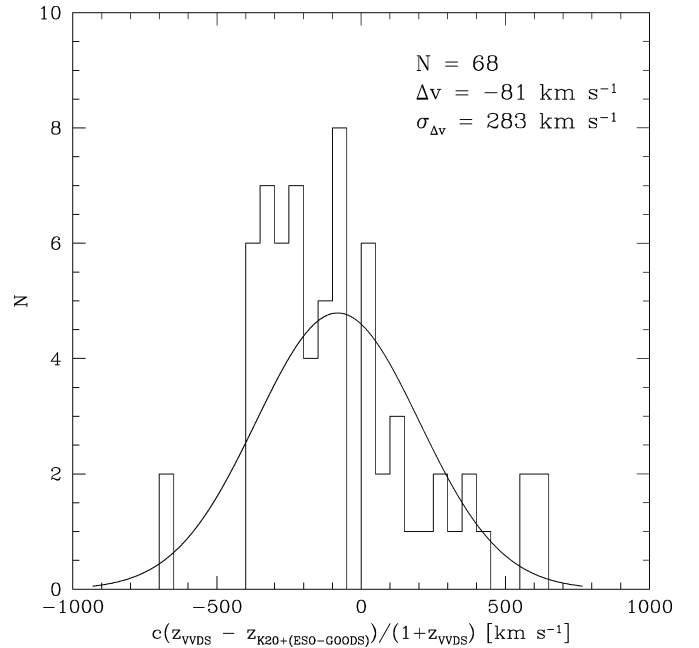


FIG. 6.—Distribution of differences in redshift measurements from a comparison between VVDS and K20 or GOODS observations.

to problems in the sky subtraction within a slit, resulting in a failure to detect a real object spectrum, and to spectra characterized by an almost absent continuum, coupled with relatively strong emission lines that escape detection. Overall, for the VVDS observations and the 16,936 slits reduced thus far, the total fraction of failures is approximately 3.9%. Eliminating the cases in which this failure is not due to the data-reduction software (as in the case of partially vignettted slits), we estimate that on average, if we exclude pure emission-line-spectra objects, VIPGI is successfully extracting a spectrum for 98% of the slits in which a real spectrum is present. For the remaining 2% of the cases, the signal is too weak for the automated object detection procedure to be successful, although the presence of an extremely faint spectrum is still discernible to the human eye.

## 7. DATA ORGANIZATION AND BROWSING TOOLS

As already discussed in § 3, an automatic reduction pipeline is not enough to guarantee a quick and timely data-reduction process for VIMOS data. An efficient organization of the data, in addition to easy procedures for data browsing, are two components that are as important as the speed of the pipeline for the global data-reduction process. It is mainly for this reason that VIPGI provides its own VIMOS-specific data organizer, which organizes the data in such a way as to prevent mistakes in the selection of the recipes' input files. Data files are placed in a rigidly predefined directory structure on the computer disk, and file names are automatically assigned to all data files, fol-

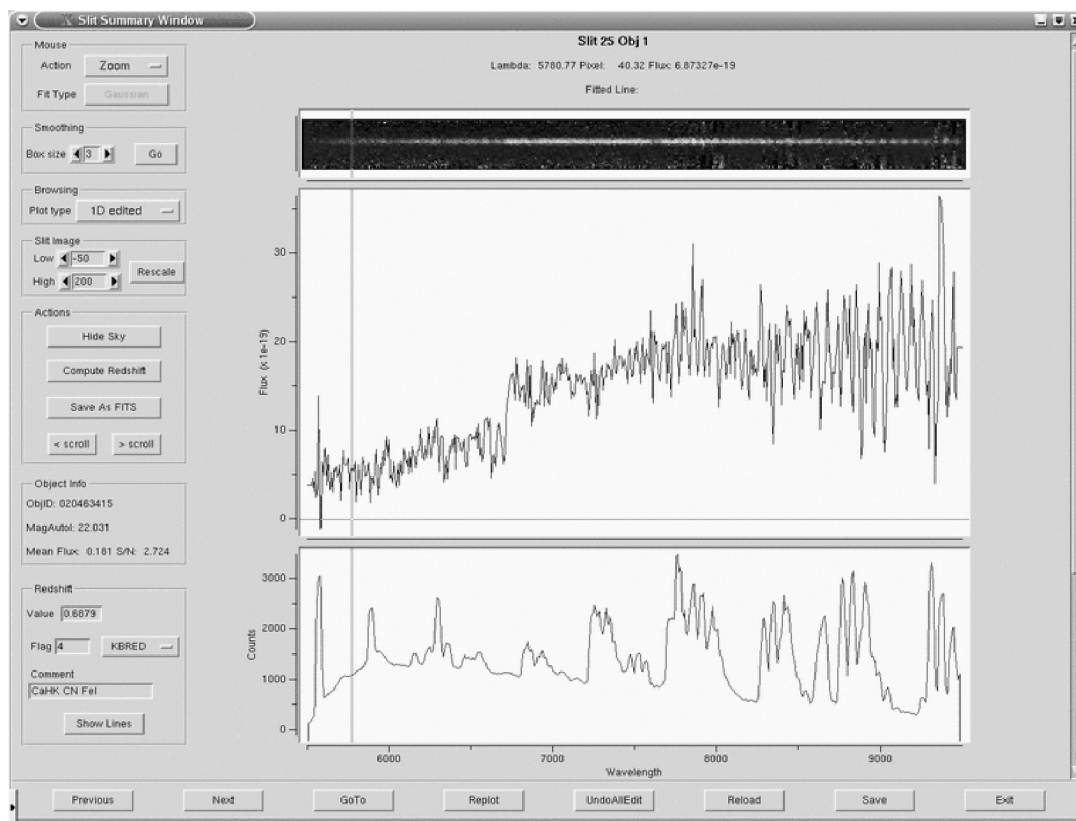


FIG. 7.—“Slit summary plot” showing the two- and one-dimensional spectra for one VVDS object, and the sky spectrum for the same slit.

lowing the organizer classification rules. Auxiliary calibration tables needed by the pipeline recipes or produced by them during the reduction process are also appended to the FITS files as extensions. In this way, the input file is one and the same for most data browsing functions, and it is left to those functions to get the correct set of data from the FITS file.

A number of different possibilities exist for browsing through one- and two-dimensionally extracted spectra. The most important tool, designed to allow the astronomer to plot and analyze MOS spectra with maximum efficiency, is the “slit summary” plotting tool shown in Figure 7. This tool allows the astronomer to plot each one-dimensional spectrum of a MOS mask, together with the corresponding two-dimensional and sky one-dimensional spectra. In this way, it is possible to visually check the reality of spectral features that are present in the one-dimensional spectrum, which could be due to sky, zeroth grating order, or fringing subtraction residuals. It is also possible to zoom in and out for all three spectra simultaneously, and to analyze the one-dimensional spectrum with a set of commands that reproduce the most important ones provided by the IRAF SPLOT task (smooth the spectrum with a simple square window function, measure its S/N over a selected wavelength interval, fit the position of a selected spectral line, and edit single-pixel or user-defined spectrum intervals). The as-

tronomer can also obtain quick redshift estimates by fitting or marking the position of a set of spectral lines and using a function that computes a list of possible redshifts based on a list of known emission and absorption lines in galaxy spectra. Once the user has chosen a possible solution, the expected positions of all the lines in the list are marked on the plot to visually inspect the goodness of the redshift determination. In addition, it is also possible to display information on the astronomical object whose spectrum is being plotted, including the wavelength calibration for the spectrum, and about the object detection process that lead to the extraction of the spectrum being analyzed. A stand-alone version of this tool is also available to the astronomical community.<sup>16</sup>

## 8. SUMMARY

We have developed VIPGI, a new tool designed to organize, reduce, and analyze data obtained with VIMOS, the imaging spectrograph built by the VIMOS Consortium for the European Southern Observatory. This tool is being used to handle all the spectroscopic data obtained for the VIMOS VLT Deep Survey, which aims to measure 150,000 galaxy redshifts.

<sup>16</sup> See <http://cosmos.mi.iasf.cnr.it/pandora/sgnaps.html>.

VIPGI provides powerful data organizing capabilities, a small set of data-reduction recipes, and dedicated data browsing and plotting tools to check the results of all critical data-reduction steps and to plot and analyze final extracted one- and two-dimensional spectra.

We have performed many data quality checks demonstrating that VIPGI can be used to achieve a data-reduction accuracy that is comparable to that obtained using standard IRAF tasks. Moreover, its high reduction speed makes our pipeline the only reasonable tool to reduce the large amount of data that a modern instrument such as VIMOS produces.

VIPGI has been extensively used in the reduction of more than 20,000 spectra for the VIMOS VLT Deep Survey, and it is now routinely being used to reduce all kinds of VIMOS data in the VIMOS Data Reduction Support Centers in Milan and Marseille.

We would like to thank Carlo Izzo, Ralf Palsa, and Paola Sartoretti for contributing to parts of the code development, Tom Osterloo for setting up the DRS coding infrastructure and the early code development, and Nicolas Devillard for his many helpful suggestions. This research has been developed within the framework of the VVDS consortium. This work has been partially supported by the CNRS-INSU and its Programme National de Cosmologie (France), and by Italian Ministry (MIUR) grants COFIN2000 (MM02037133) and COFIN2003 (2003020150). The VLT-VIMOS observations have been carried out on guaranteed time (GTO) allocated by the European Southern Observatory (ESO) to the VIMOS Consortium, under a contractual agreement between the Centre National de la Recherche Scientifique of France, heading a consortium of French and Italian institutes, and ESO, to design, manufacture, and test the VIMOS instrument.

## REFERENCES

- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393  
 Bottini, D., et al. 2005, *PASP*, 117, 996  
 Colless, M., Ellis, R. S., Taylor, K., & Hook, R. N. 1990, *MNRAS*, 244, 408  
 Colless, M., et al. 2001, *MNRAS*, 328, 1039  
 Greenfield, P., & White, R. L. 2000, in *ASP Conf. Ser. 216, Astronomical Data Analysis Software and Systems IX*, ed. N. Manset, C. Veillet, & D. Crabtree (San Francisco: ASP), 59  
 Horne, K. 1986, *PASP*, 98, 609  
 Le Fèvre, O., Crampton, D., Lilly, S. J., Hammer, F., & Tresse, L. 1995, *ApJ*, 455, 60  
 Le Fèvre, O., et al. 2000, *Proc. SPIE*, 4008, 546  
 ———. 2002, *Messenger*, 109, 21  
 ———. 2005, *A&A*, 439, 845  
 Mink, D. J. 1998, *ASP Conf. Ser. 172, Astronomical Data Analysis Software and Systems VII*, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 498  
 Paiono, L., Garilli, B., Scoddeggio, M., & Franzetti, P. 2004, in *Conf. Proc., Astronomical Data Analysis III (astro-ph/0406117)*  
 Pence, W. D. 1998, *ASP Conf. Ser. 172, Astronomical Data Analysis Software and Systems VII*, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 487  
 Scoddeggio, M., Zanichelli, A., Garilli, B., Le Fèvre, O., & Vettolani, G. 2000, *ASP Conf. Ser. 283, Astronomical Data Analysis Software and Systems X*, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne (San Francisco: ASP), 451  
 Stoughton, C., et al. 2002, *Proc. SPIE*, 4836, 339  
 Strauss, M. A., et al. 2002, *AJ*, 124, 1810  
 Zanichelli, A., et al. 2005, *PASP*, 117, 1271